On combining information-theoretic and cryptographic approaches to network coding security against the pollution attack

Svitlana Vyetrenko
California Institute of Technology
Pasadena, CA 91125, USA
Email: svitlana@caltech.edu

Aditya Khosla
Stanford University
Stanford, CA 94305, USA
Email: aditya86@stanford.edu

Tracey Ho
California Institute of Technology
Pasadena, CA 91125, USA
Email: tho@caltech.edu

***Abstract*—In this paper we consider the pollution attack in network coded systems where network nodes are computationally limited. We consider the combined use of cryptographic signature based security and information theoretic network error correction and propose a fountain-like network error correction code construction suitable for this purpose.**

I. Introduction

In this paper we consider the problem of adversarial errors in single-source multicast networks with limited computational power (e.g. wireless or sensor networks). Most existing results on information theoretic multicast network error correction assume a given bound on the number of adversarial errors, e.g. [5], [8], for which random linear network coding achieves capacity [9]. If z_u is the upper bound on the number of errors that can occur in the network, noncoherent network coding is used at all nodes and M is the minimum cut of the network, the error correcting code that achieves information rate $M - 2z_u$ can be constructed [5].

An alternative approach to network error correction is equipping each network packet with a cryptographic signature, e.g. [10], [1]. Then, if each network node checks all packets and all nodes perform network coding, for any number of network errors z the information rate $M - z$ can be achieved without the need for further information-theoretic error correction. However, performing signature checks at all network nodes may limit throughput in a network with limited computational resources, since such cryptographic operations are typically more expensive than network coding operations. Therefore, we are interested in combining benefits of both approaches. We consider probabilistic verification of a subset of packets in conjunction with information-theoretic redundancy so as to achieve intermediate information rates r with

$$M - 2z_u \leq r \leq M - z$$

subject to computational budget constraints at each node.

[0]This work was partly funded by subcontract 069153 issued by BAE Systems National Security Solutions, Inc. and supported by the Defense Advanced Research Projects Agency (DARPA) and the Space and Naval Warfare System Center (SPAWARSYSCEN), San Diego under Contract No. N66001-08-C-2013, and by Caltech's Lee Center for Advanced Networking. Part of this work was done while Aditya Khosla was at Caltech.

In order to solve this problem, we need to develop a framework to use network error correction in a probabilistic setting. In existing network error correcting algorithms, the deterministic bound on the number of erroneous packets needs to be known in advance for code construction [5]. This can result in a very conservative upper bound when packets are checked probabilistically. In this paper we propose a fountain-like network error correcting code construction that can be used in networks where the upper bound on the number of errors is unknown a priori. Instead of including a fixed number of redundant bits in each packet, we incrementally add redundancy until decoding succeeds.

II. Problem statement

Let $\mathcal{G}$ be an acyclic network with source $\mathcal{S}$ and sink $\mathcal{T}$. Let M be the minimum cut of $\mathcal{G}$. The nodes of $\mathcal{G}$ are limited in computational power and outgoing capacity. Let n be the number of nodes in $\mathcal{G}$. Errors can occur on some links of $\mathcal{G}$.

Let N_{in}^i be the number of packets incoming to node i, and let N_{out}^i be the number of packets outgoing from node i. Let A_i be the computational budget available at node i. Given A_i, we assume that in addition to forwarding all outgoing packets, each node i has the capacity to check a fraction ρ_i of incoming packets and to form a fraction γ_i of outgoing packets by creating random linear combinations of packets incoming to node i, so that

$$\rho_i N_{in}^i + \gamma_i N_{out}^i \leq A_i.$$

Let $\overrightarrow{\rho} = (\rho_1, \rho_2, \ldots, \rho_n)$ be the vector that defines the checking strategy at nodes of $\mathcal{G}$. Let $\overrightarrow{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ be the vector that defines the network coding strategy at nodes of $\mathcal{G}$. Let $\overrightarrow{A} = (A_1, A_2, \ldots, A_n)$ be the vector of computational budgets available at nodes of $\mathcal{G}$. Let

$$\Sigma = \{\overrightarrow{\rho}, \overrightarrow{\gamma} \mid \overrightarrow{\rho}, \overrightarrow{\gamma} \text{ are feasible for a given } \overrightarrow{A}\}$$

be the set of all strategies feasible at nodes of $\mathcal{G}$ for a given budget constraint $\overrightarrow{A}$. Let $r_\sigma(\overrightarrow{A})$ be the information rate that can be achieved for a given $\sigma \in \Sigma$ and $\overrightarrow{A}$.

In this paper, we focus on how to construct the error correcting code that achieves $r_\sigma(\overrightarrow{A})$ for a given $\sigma \in \Sigma$. For

each $\sigma \in \Sigma$ the number of erroneous packets available at the sink is unknown in advance, therefore, we want to construct the code that can adapt to the actual number of errors present at the sink. Moreover, if an erroneous packet injected to link l remains unchecked due to computational budget constraint and random linear coding is performed, any subsequent signature check will identify packets contained on links downstream of l as erroneous and will eliminate them. Therefore, we require that the code that we construct be applicable in networks with unknown time-varying minimum cut and number of errors.

III. Code construction

Throughout the paper, we use the following notation. For any matrix A, let rows(A) denote the set of vectors that form rows of A. Let I_a denote an $a \times a$ identity matrix. Also, let $\mathbf{i_a}$ denote an $a^2 \times 1$ vector that is obtained by stacking columns of I_a one after the other. Let $\mathbb{F}_q$ be the finite field over which coding occurs. Each source packet contains K symbols from $\mathbb{F}_q$.

A. Encoder

In each block $\mathcal{S}$ transmits BK independent information symbols from $\mathbb{F}_q$ to $\mathcal{T}$. Let W be a $B \times K$ matrix whose elements are the information symbols. The source transmits rows(X_0), where $X_0 = \begin{pmatrix} W & I_B \end{pmatrix}$. Suppose that while transmitting rows(X_0) by means of random linear network coding, the network has incurred $z_0 > 0$ errors. Then since there are z_0 additions and $d_0 = B - z_0$ deletions to/from rowspace(X_0), $\mathcal{T}$ would not be able to recover X_0.

By [2], if there are $d_0 = B - z_0$ deletions and no additions from rowspace(X_0), sending $\delta = d_0$ additional linear combinations of rows(X_0) ensures successful decoding. Similarly, by [5], in case of z_0 additions and no deletions, sending $\sigma K > z_0 K$ redundant bits helps to decode. By making use of the two above-mentioned ideas, we propose an iterative algorithm that resembles a "digital error fountain" by incrementally adding redundancy, that ensures decoding of the source packets in finite number of iterations.

An end to end error detection scheme is needed so that the sink can determine when decoding is successful. For instance, the source can include a cryptographic signature, e.g. [1], in each packet. Upon failing to decode X_0 successfully from the initial transmission, $\mathcal{S}$ sends an additional batch of σ_1 linearly independent redundant packets and δ_1 linearly dependent redundant packets, and $\mathcal{T}$ attempts to decode using both the initial and the redundancy batch. Additional batches of redundant symbols are transmitted until decoding succeeds, whereupon the sink sends feedback telling the source to move onto the next batch.

The ith stage of the reencoding algorithm can be generalized as follows (see Fig. 1):

- Let $\sigma_i = m/2$. The encoder arranges the matrix of information symbols W in an $BK \times 1$ vector $\mathbf{w}$. Let S_i be a $\sigma_i K \times BK$. Define a vector of redundant symbols $\mathbf{y_i}$ as

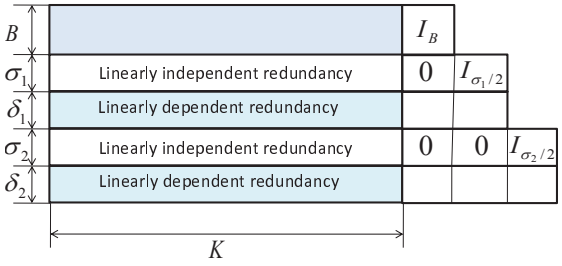


Fig. 1. Code construction.

$$\mathbf{y_i} = S_i \mathbf{w} \text{ or, equivalently,}$$
$$\begin{pmatrix} S_i & -I_{\sigma_i K} \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ \mathbf{y_i} \end{pmatrix} = 0. \quad (1)$$

After computing $\mathbf{y_i}$, the encoder arranges it into a $\sigma_i \times (K + (i+1)m)$ matrix Y_i column by column. Set

$$A_i^1 = \begin{pmatrix} Y_i^1 & 0 & I_{\sigma_i} \end{pmatrix}, \quad (2)$$

where 0 is a $\sigma_i \times (B + (i-1)m)$ matrix with zero entries.

- Let $\delta_i = m/2$. Let D_i be a $\delta_i \times \left(B + \sum_{j=1}^{i} \sigma_j \right)$ matrix with random entries from $\mathbb{F}_q$. Define a $\delta_i \times (K + (B + im))$ matrix A_i^2 as

$$A_i^2 = D_i \begin{pmatrix} X_0 & 0 & 0 & \dots & 0 \\ \hline A_1^1 & & 0 & \dots & 0 \\ \hline & A_2^1 & & \dots & 0 \\ \hline & & \dots & & \\ \hline & & & A_i^1 & \end{pmatrix}. \quad (3)$$

- At the ith stage, the source transmits $X_i = \begin{pmatrix} A_i^1 \\ A_i^2 \end{pmatrix}$.

B. Decoder

Let z_i be the number of errors, i.e. packets corrupted by the adversary, at the ith stage. Let Z_i be the matrix whose rows are the error packets injected to the network at the ith stage that are linearly independent of the X_i packets, i.e. rowspace$(X_i) \cap$ rowspace$(Z_i) = 0$. Let

$$Y_i = T_i X_i + Q_i Z_i \quad (4)$$

be the matrix, such that rows(Y_i) are the packets received at $\mathcal{T}$ at the ith stage, where T_i is the transfer matrix from all links in $\mathcal{G}$ to the packets received at $\mathcal{T}$, and Q_i is the transfer matrix from error packets to the packets received at $\mathcal{T}$ at the ith stage. For notational convenience, define

$$
\begin{aligned}
Y^i &= \begin{pmatrix} Y_0 \\ Y_1 \\ \dots \\ Y_i \end{pmatrix} \quad T^i = \begin{pmatrix} T_0 & 0 & \dots & 0 \\ 0 & T_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & T_i \end{pmatrix} \\
Q^i &= \begin{pmatrix} Q_0 & 0 & \dots & 0 \\ 0 & Q_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & Q_i \end{pmatrix} \\
X^i &= \begin{pmatrix} X_0 & 0 & 0 & \dots & 0 \\ \hline X_1 & 0 & \dots & 0 \\ \hline X_2 & & \dots & 0 \\ \hline \multicolumn{4}{c}{\dots} \\ \hline \multicolumn{4}{c}{X_i} \end{pmatrix} \\
Z^i &= \begin{pmatrix} Z_0 & 0 & 0 & \dots & 0 \\ \hline Z_1 & 0 & \dots & 0 \\ \hline Z_2 & & \dots & 0 \\ \hline \multicolumn{4}{c}{\dots} \\ \hline \multicolumn{4}{c}{Z_i} \end{pmatrix}
\end{aligned}
$$

Note that for any i we can write

$$Y^i = T^i X^i + Q^i Z^i.$$

The source transmits at the minimum cut rate M. Thus, X_0 is transmitted in $N_B = \frac{B}{M}$ time units and each X_i, $i = 1, 2, \dots$ is transmitted in $N_m = \frac{m}{M}$ time units. For each $j = 1, 2, \dots, B$ denote the part of X_0 transmitted at the jth time unit by X_0^j. Similarly, for each $j = 1, 2, \dots, m$, $i = 1, 2, \dots$ denote the part of X_i by transmitted at the jth time unit by X_i^j. For each i, j, define E_i^j to be a random variable that corresponds to the number of errors that occurred in $\mathcal{G}$ while transmitting X_i^j. Define $E_0 = \sum_{j=1}^{N_B} E_i^j$ and $E_i = \sum_{j=1}^{N_m} E_i^j$, $i = 1, 2, \dots$. Recall that $\sigma_i = \delta_i = \frac{m}{2}$.

Lemma 1: Suppose that for each i, j, there exists $\epsilon_i^j > 0$ such that

$$\mathbb{E}[E_i^j] < \frac{M}{2} - \epsilon_i^j. \tag{5}$$

Then for some finite N, we will have

$$\sum_{i=0}^{N} z_i < \sum_{i=1}^{N} \delta_i \tag{6}$$

$$\sum_{i=0}^{N} z_i < \sum_{i=1}^{N} \sigma_i \tag{7}$$

Proof: Let $\epsilon = \min_{i,j} \epsilon_i^j$. Note that

$$
\begin{aligned}
\mathbb{E}[E_0] &= \sum_{j=1}^{N_B} \mathbb{E}[E_0^j] < \frac{B}{2} - \epsilon N_B < \frac{B}{2} \\
\mathbb{E}[E_i] &= \sum_{j=1}^{N_m} \mathbb{E}[E_i^j] < \frac{m}{2} - \epsilon N_m, i = 1, 2, \dots
\end{aligned}
$$

Then for $L^* > \frac{B}{2\epsilon N_m}$

$$
\begin{aligned}
\sum_{i=0}^{L^*} \mathbb{E}[E_i] &< \mathbb{E}[E_0] + \frac{mL^*}{2} - L^* \epsilon N_m \\
&< \frac{B}{2} + \frac{mL^*}{2} - L^* \epsilon N_m < \frac{mL^*}{2}.
\end{aligned}
$$

Therefore, for some finite $N > L^*$, we will have

$$\sum_{i=0}^{N} z_i \leq \sum_{i=0}^{N} \mathbb{E}[E_i] < \frac{mN}{2}, \tag{8}$$

hence, we have $\sum_{i=0}^{N} z_i < \sum_{i=1}^{N} \delta_i$ and $\sum_{i=0}^{N} z_i < \sum_{i=1}^{N} \sigma_i$. ■

Lemma 2: If

$$\sum_{i=0}^{N} z_i \leq \sum_{i=1}^{N} \delta_i, \tag{9}$$

then with high probability columns of T^N and Q^N span disjoint vector spaces.

Proof: Note that $\sum_{i=1}^{N} \delta_i + \sum_{i=1}^{N} \sigma_i + B = Nm + B$. Then by adding $\sum_{i=1}^{N} \sigma_i + B$ to both sides of (9), we get

$$\sum_{i=0}^{N} z_i + \sum_{i=1}^{N} \sigma_i + B \leq Nm + B,$$

or

$$\text{rank}(X^N) + \text{rank}(Z^N) \leq Nm + B.$$

Therefore, if the error packets were replaced by additional source packets, the total number of source packets would be at most $Nm + B$. By [3], with high probability, random linear network coding allows $\mathcal{T}$ to decode all source packets. This corresponds to $\begin{pmatrix} T^N & Q^N \end{pmatrix}$ having full column rank, hence, column spaces of T^N and Q^N being disjoint except in the zero vector. ■

Let N be such that conditions (6)-(7) are satisfied. Then in order to decode, we need to solve the following system of linear equations:

$$Y^N = T^N X^N + Q^N Z^N \tag{10}$$

$$\begin{pmatrix} S_1 & -I_{\frac{mK}{2}} & \dots & 0 \\ S_2 & 0 & \dots & 0 \\ & & \dots & \\ S_N & 0 & \dots & -I_{\frac{mK}{2}} \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ \mathbf{y_1} \\ \dots \\ \mathbf{y_N} \end{pmatrix} = 0 \tag{11}$$

Theorem 1: Let N be such that equations (6) and (7) are satisfied. Then with probability greater than $1 - q^{-\epsilon K}$, the system of linear equations (10)-(11) can be solved for $\mathbf{x}$.

Proof: The proof of this theorem is constructive and is similar to [5]. Note that

$$X^N = \left(\begin{array}{ccccc} X_0 & 0 & 0 & \dots & 0 \\ \hline X_1 & 0 & \dots & 0 \\ \hline X_2 & & \dots & 0 \\ \hline \dots \\ \hline X_N \end{array}\right) = \left(\begin{array}{ccccc} X_0 & 0 & 0 & \dots & 0 \\ \hline A_1^1 & 0 & \dots & 0 \\ A_1^2 & 0 & \dots & 0 \\ \hline A_2^1 & & \dots & 0 \\ A_2^2 & & \dots & 0 \\ \hline \dots \\ \hline A_N^1 \\ A_N^2 \end{array}\right).$$

Define

$$\begin{aligned} X &= \left(\begin{array}{ccccc} X_0 & 0 & 0 & \dots & 0 \\ \hline A_1^1 & 0 & \dots & 0 \\ \hline A_2^1 & & \dots & 0 \\ \hline \dots \\ \hline A_N^1 \end{array}\right) \\ &= \left(\begin{array}{ccccc} W & I_B & 0 & \dots & 0 \\ Y_1 & 0 & I_{m/2} & \dots & 0 \\ & & \dots & & \\ Y_N & 0 & 0 & \dots & I_{m/2} \end{array}\right) \end{aligned} \quad (12)$$

Let $0_{a,b}$ denote a zero matrix with a rows and b columns. Note that by (3) $X^N = D^N X$, where

$$D^N = \left(\begin{array}{cccccc} I_B & 0_{B,m/2} & 0_{B,m/2} & \dots & 0_{B,m/2} \\ \hline 0_{m/2,B} & I_{m/2} & 0_{m/2,m/2} & \dots & 0_{m/2,m/2} \\ \hline D_1 & & 0_{m/2,m/2} & \dots & 0_{m/2,m/2} \\ \hline 0_{m/2,B} & I_{m/2} & 0_{m/2,m/2} & \dots & 0_{m/2,m/2} \\ \hline D_2 & & & \dots & 0_{m/2,m/2} \\ \hline \dots & \dots & \dots & \dots & \dots \\ \hline 0_{m/2,B} & 0_{m/2,m/2} & 0_{m/2,m/2} & \dots & I_{m/2} \\ \hline D_N \end{array}\right)$$

Let $T = T^N D^N$. Then (10) is equivalent to

$$Y = TX + QZ, \quad (13)$$

where $Y = Y^N$, $Q = Q^N$ and $Z = Z^N$.

Let $b = B + \sum_{i=1}^{N} \sigma_i = B + \frac{mN}{2}$. The identity matrix of dimension b sent by $\mathcal{S}$ undergoes the same transformation as the rest of the batch. Hence, $\hat{T} = TI_b + QL$, where $\hat{T}$ and L are the columns that correspond to the location of the identity matrix in Y and Z respectively. Then we can write

$$Y = \hat{T}X + Q(Z - LX) = \hat{T}X + E,$$

with $E = Q(Z - LX)$.

Assume that Y full row rank, otherwise, discard linearly dependent rows of Y. Define $z = \text{rank}(QZ)$. By Lemma 2 $z = \text{rank}(Y) - b$ and T^N and Q span disjoint vector spaces. Since columns of $T = T^N D^N$ are linear combinations of columns of T^N, T and Q also span disjoint vector spaces. Because the decoder cannot directly estimate the basis for the column space of E, it instead chooses a proxy error matrix T'' whose columns act as a proxy error basis for the columns of E. T'' is chosen as the matrix that corresponds to the first z columns of Y. As in [5], we then have

$$Y = \left(\begin{array}{cc} T'' & \hat{T} \end{array}\right) \left(\begin{array}{ccc} I_z & F^Z & 0 \\ 0 & F^X & I_b \end{array}\right). \quad (14)$$

Let $X = \left(\begin{array}{ccc} J_1 & J_2 & J_3 \end{array}\right)$, where J_1 corresponds to the first z columns of X, J_3 corresponds to the last b columns of X, and J_2 corresponds to the remaining columns of X. Then by Claim 4 in [5], (14) is equivalent to the matrix equation

$$\hat{T}J_2 = \hat{T}(F^X + J_1 F^Z). \quad (15)$$

Now, in order to decode, we need to solve the system formed by the linear equations (11) and (15).

For $i = 1, 2$ denote by $\mathbf{j_i}$ the vector obtained by stacking the columns of J_i one on top of the other. Note that by (12),

$$\left(\begin{array}{c} \mathbf{j_1} \\ \mathbf{j_2} \end{array}\right) = P \left(\begin{array}{c} \mathbf{w} \\ \mathbf{y_1} \\ \dots \\ \mathbf{y_N} \end{array}\right),$$

where P is a permutation matrix.

Denote by $\mathbf{f}^X$ the vector formed by stacking columns of the matrix F^X one on top of another, and by $f_{i,j}$ the (i,j)th entry of the matrix F^Z. Let $\alpha = K - z$. The system of linear equations given by (11) and (15) can be written in matrix form as

$$A \left(\begin{array}{c} \mathbf{j_1} \\ \mathbf{j_2} \end{array}\right) = \left(\begin{array}{c} \hat{T}\mathbf{f}^X \\ \mathbf{0} \\ \dots \\ \mathbf{0} \end{array}\right),$$

where A is given by

$$A = \left(\begin{array}{cccc|cccc} -f_{1,1}\hat{T} & -f_{2,1}\hat{T} & \dots & -f_{z,1}\hat{T} & \hat{T} & 0 & \dots & 0 \\ -f_{1,2}\hat{T} & -f_{2,2}\hat{T} & \dots & -f_{z,2}\hat{T} & 0 & \hat{T} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -f_{1,\alpha}\hat{T} & -f_{2,\alpha}\hat{T} & \dots & -f_{z,\alpha}\hat{T} & 0 & 0 & \dots & \hat{T} \\ \hline S_P \end{array}\right)$$

with $S_P = \left(\begin{array}{ccccc} S_1 & -I_{\frac{mK}{2}} & 0 & \dots & 0 \\ S_2 & 0 & -I_{\frac{mK}{2}} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ S_N & 0 & 0 & \dots & -I_{\frac{mK}{2}} \end{array}\right) P^{-1}$.

In order to show that we can decode, we need to prove that A has full column rank. By Lemma 2, $\hat{T}$ is a $(b+z) \times b$ matrix of full column rank. Therefore, the last αb columns of A have full column rank. Denote the first z block-columns of A by $\{u_1, u_2, \dots, u_z\}$, and the last α block-columns of A by $\{v_1, v_2, \dots, v_\alpha\}$. For each i, let $u_i = \left(\begin{array}{cc} u_i^1 & u_i^2 \end{array}\right)^T$, where u_i^1 are the first $\alpha(b+z)$ rows and u_i^2 are the remaining rows of u_i. Similarly, let $v_i = \left(\begin{array}{cc} v_i^1 & v_i^2 \end{array}\right)^T$, where v_i^1 are the first $\alpha(b+z)$ rows and v_i^2 are the remaining rows of v_i. Note that for each $i = 1 \dots z$, $u_i^1 + \sum_j f_{i,j} v_i^1 = 0$. Define $w_i = u_i^2 + \sum_j f_{i,j} v_i^2$. Let $\tilde{A}$ be the resulting matrix after Gaussian elimination is performed on the upper left-hand side of A. A has full rank iff the lower submatrix of $\tilde{A}$ formed by w_i and v_i^2 has full rank. Note that since P is a permutation matrix, P^{-1} is also a permutation matrix. Therefore, S_P is a permutation of columns of the random matrix $S = \left(\begin{array}{c} S_1 \\ S_2 \\ \dots \\ S_N \end{array}\right)$

and the identity matrix; hence, u_i^2 and v_i^2 are the columns of S and the identity matrix. Since entries of S are independently and uniformly distributed in $\mathbb{F}_q$, so are w_i for fixed values of $f_{i,j}$. The probability that A does not have full column rank is $1 - \prod_{l=1}^{bz} \left(1 - \frac{1}{q^{\sum \sigma_i K - l + 1}}\right)$, which is upper-bounded by $q^{bz - \sum \sigma_i K}$. By the union bound over all $q^{\alpha z}$ possible values of variables $f_{i,j}$, we have $q^{bz - \sum \sigma_i K + \alpha z} \leq q^{K(z - \sum \sigma_i)}$. Therefore, decoding succeeds with probability at least $q^{-K\epsilon}$ if $\sum \sigma_i > z + \epsilon$, which follows from equation (7). ■

Theorem 2: For each i, j, let E_i^j be random variables with the same mean such that (5) is satisfied. Let N be such that equations (6)-(7) are satisfied. Then the above-described code construction achieves the information rate

$$r \leq M_A - 2\mathbb{E}[E_0^1] - \epsilon, \tag{16}$$

where M_A is the average throughput of linearly independent packets, and ϵ decreases with increasing B.

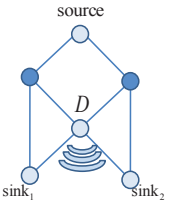
Proof: Define $\epsilon_1 = \frac{mN}{2} - \sum_{i=0}^{N} \mathbb{E}[E_i]$. By (8) $\epsilon_1 > 0$. Since for each i,j, the actual minimum cut of the network varies depending on the strategy used, define M_i^j to be the throughput of linearly independent packets while transmitting X_i^j. Then the achievable rate is given by:

$$\begin{aligned} r &\leq \frac{\sum_{j=1}^{N_B} M_0^j + \sum_{i=1}^{N}\sum_{j=1}^{N_m} M_i^j - \sum_{i=1}^{N}(\sigma_i + \delta_i)}{N_B + NN_m} \\ &= M_A - 2\frac{\sum_{j=1}^{N_B}\mathbb{E}[E_0^j] + \sum_{i=1}^{N}\sum_{j=1}^{N_m}\mathbb{E}[E_i^j]}{N_B + NN_m} - \frac{2\epsilon_1}{N_B + NN_m} \\ &= M_A - 2\frac{\mathbb{E}[E_0^1](N_B + NN_m)}{N_B + NN_m} - \frac{2\epsilon_1 M}{B + mN}, \end{aligned}$$

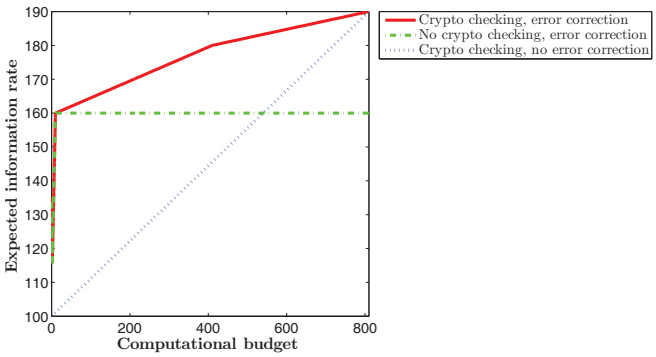
where $\epsilon = \frac{2\epsilon_1 M}{B+mN}$. ■

IV. Example: Wireless Butterfly Network

We consider a wireless butterfly network where a computationally limited network coding node D receives z adversarial packets (see Fig. 2(a)). For a varying computational budget constraint, we compare three strategies: when network error correction is performed without cryptographic checking, when cryptographic checking is performed without network error correction, and when both cryptographic checking and network error correction are performed. We derived analytical expressions for the expected information rate for all three strategies, which are plotted in Fig. 2(b). Note that using our code construction the expected information rate can be approached.



(a) Wireless butterfly network topology



(b) Expected information rate for network of Fig. 2(a) with equal capacity links, minimum cut $= 200$, $z = 20$, $\frac{\text{cost of coding}}{\text{cost of checking}} = 40$.

Fig. 2. Example: wireless butterfly network example

Acknowledgment

We thank Ted Dikaliotis for his invaluable help in preparing this paper.

References

[1] D.Boneh, D.Freeman, J.Katz, B.Waters, "Signing a linear subspace: signature schemes for network coding", *Lecture Notes Comp. Science*, vol. 5443, pp. 68-87, 2009.

[2] A.F.Dana, R.Gowaikar, R.Palanki, B.Hassibi, M.Effros, "Capacity of wireless erasure networks", *IEEE Trans. Inf. Theory*, vol. 52, pp. 789-804, 2006.

[3] T.Ho, R.Koetter, M.Medard, M.Effros, J.Shi, D.Karger, "A random linear network coding approach to multicast", *IEEE Trans. Inf. Theory*, Vol. 52, No. 10, pp.4413-4430, 0ct. 2006

[4] T.Ho, B.Leong, R.Koetter, M.Medard, M.Effros, D.Karger, "Byzantine modifiation detection in muticast networks with random network coding", *IEEE Trans. Inf. Theory, Special Issue on Inf. Theor. Security*, Vol. 54, No. 6, pp. 2798-2803, Jun. 2008.

[5] S.Jaggi, M.Langberg, S.Katti, T.Ho, D.Katabi, M.Medard, "Resilient network coding in the presence of Byzantine adversaries", *IEEE Trans. Inf. Theory, Special Issue on Inf. Theor. Security*, Vol. 54, No. 6, pp. 2596-2603, Jun. 2008.

[6] R.Koetter, F.Kschischang, "Coding for the errors and erasures in random network coding", *IEEE Trans. Inf. Theory*, vol. 54, pp. 3579-3591, Aug. 2008.

[7] D.Silva, F.R.Kschischang, "On metrics for error correction in network coding", submitted to *IEEE Trans. Inform. Theory*, 2008.

[8] S.Yang, R.Yeung, "Refined coding bounds for network error correction", *IEEE ITW on Inf. Theory for Wireless Netw.*, Jul. 2007.

[9] R.Yeung, N.Cai, "Network error correction, Part I: Basic concepts and upper bounds", *Commun. Inf. Syst.*, Vol. 6, No. 1. pp. 19-36, 2006.

[10] F.Zhao, T.Kalker, M.Medard, K.J.Han, "Signatures for content distribution with network coding", *ISIT*, Jun. 2007.